POSITIVE DEFINITE METRIC SPACES

MARK W. MECKES

Abstract. Magnitude is a numerical invariant of finite metric spaces, recently introduced by T. Leinster, which is analogous in precise senses to the cardinality of finite sets or the Euler characteristic of topological spaces. It has been extended to infinite metric spaces in several a priori distinct ways. This paper develops the theory of a class of metric spaces, positive definite metric spaces, for which magnitude is more tractable than in general. Positive definiteness is a generalization of the classical property of negative type for a metric space, which is known to hold for many interesting classes of spaces. It is proved that all the proposed definitions of magnitude coincide for compact positive definite metric spaces and further results are proved about the behavior of magnitude as a function of such spaces. Finally, some facts about the magnitude of compact subsets of ℓ_p^n for $p \le 2$ are proved, generalizing results of Leinster for $p = 1, 2$, using properties of these spaces which are somewhat stronger than positive definiteness.

1. Introduction

Magnitude is a canonical numerical invariant of finite metric spaces recently introduced by Tom Leinster [22, 24], motivated by considerations from category theory. The same notion appeared earlier, although it was not really developed, in connection with measuring biodiversity [34]. Magnitude is analogous in a precise sense to the Euler characteristic of topological spaces or partially ordered sets, and to the cardinality of finite sets, and it may be interpreted as the *effective number* of points of a space. The definition of magnitude was extended to infinite metric spaces in various ways in the papers [25, 41, 42, 24]. In this setting, magnitude turns out to have close connections (some proved, and some only conjectural at present) to classical invariants of geometric measure theory and integral geometry, including Hausdorff dimension and intrinsic volumes of convex bodies and Riemannian manifolds. This paper is devoted to developing the theory of a particular class of metric spaces, *positive definite* metric spaces, for which the theory of magnitude is more tractable than in general. Examples of positive definite metric spaces include many spaces of interest, including all subsets of L_p when $1 \le p \le 2$, round spheres, and hyperbolic spaces.

Given a finite metric space (A, d), its **similarity matrix** is the matrix $\zeta_A \in \mathbb{R}^{A\times A}$ given by $\zeta_A(x, y) = e^{-d(x,y)}$. A **weighting** for A is a vector $w \in \mathbb{R}^A$ such that $\zeta_A w = \mathbf{1}$, the vector indexed by A whose entries are all 1; i.e., $\sum_{y\in A} e^{-d(x,y)}w(y) = 1$ for every $x \in A$. If a weighting w for A exists, then the **magnitude** of A is defined to be $|A| = \sum_{x\in A} w(x)$. (It is easy to check that if multiple weightings for A exist, they give the same value for $|A|$.) The reader is referred to [24] for the category-theoretic motivation of this definition, and to [24, 25, 41, 42] for discussions of various intuitive interpretations of magnitude.

As a function of an arbitrary finite metric space, magnitude may exhibit a number of pathological behaviors, the most obvious of which is that it may be undefined. One simple condition that prevents this unpleasant situation (as well as other pathologies; see [24,

Research supported in part by NSF grant DMS-0902203.

Section 2.4] for a number of relevant results) is if the similarity matrix ζ_A is positive definite. In that case ζ_A is in particular invertible, and $w = \zeta_A^{-1}\mathbf{1}$ is a weighting for A. A finite metric space A is called **positive definite** (respectively, **positive semidefinite**) if ζ_A is a positive definite (positive semidefinite) matrix. Besides the fact that magnitude is always defined, other nice properties of the class of positive definite finite metric spaces include that magnitude is positive and monotone. That is, if A is a positive definite space and $\emptyset \neq B \subseteq A$ (so that B is positive definite as well), then $0 < |B| \leq |A|$.

Three different approaches to extending the definition of magnitude to infinite spaces were taken in [25, 42, 24]. One of the purposes of this paper is to show that these approaches are essentially equivalent in the presence of an appropriate positive definiteness assumption. To that end, an arbitrary metric space A is defined to be **positive definite** (respectively **positive semidefinite**) if each of its finite subsets is positive definite (positive semidefinite) with their induced metrics. Other aims of this paper are to investigate the regularity of magnitude as a function of a positive definite metric space, and to clarify somewhat which metric spaces are and are not positive definite.

As will be seen in Section 3, a natural strengthening of positive definiteness is equivalent to the classical property of negative type for metric spaces. Although the terminology is more recent, negative type was introduced and studied by Menger [29] and Schoenberg [31, 33], and is well-studied in the literature on metric embeddings; see e.g. [39, 3]. Thus the theory of magnitude naturally leads back to this classical notion.

For clarity, a **metric space** $A = (A, d)$ here consists of a nonempty set A equipped with a **metric** $d\colon A \times A \to [0, \infty)$ such that

- $d(x, y) = 0$ if and only if $x = y$,
- $d(x, y) = d(y, x)$ for every $x, y \in A$, and
- $d(x, y) \leq d(x, z) + d(z, x)$ for every $x, y, z \in A$.

The category-theoretic motivation for the definition of magnitude is based in part on the observation by Lawvere [21] (which will not be explained here) that a metric space is a particular instance of an *enriched category*. As pointed out to the author by T. Leinster, of the properties of d above only the triangle inequality and the fact that $d(x, x) = 0$ for every x are necessary to Lawvere's observation (which moreover even allows infinite distances); whereas some classical results used in this paper, for example [33, Theorem 1], require the symmetry property of d but not the triangle inequality. Attention will therefore be restricted to the classical definition of a metric space as given above.

The rest of this paper is organized as follows. The remainder of this section establishes some additional notation and terminology. Section 2 shows the equivalence, for compact positive definite metric spaces, of several proposed definitions of magnitude, and investigates continuity properties of magnitude as a function of the metric space. Section 3 discusses sufficient conditions for positive definiteness, in particular showing the connection with negative type, and presents some counterexamples. Finally, Section 4 generalizes some results of Leinster [24] about the magnitude of subsets of Euclidean space ℓ_2^n and taxicab space ℓ_1^n to ℓ_p^n spaces for $p < 2$, using properties of those spaces which are stronger than positive definiteness.

Notation, terminology, and conventions

It will be useful to consider two general simple transformations of a metric on a fixed set. If $A = (A, d)$ is a metric space, $t \in (0, \infty)$, and $\alpha \in (0, 1]$, then tA is shorthand for the metric space (A, td) and A^α is shorthand for the metric space (A, d^α).

For a metric space A, $M(A)$ denotes the space of finite signed Borel measures on A. Unless otherwise specified, a **measure** will always refer to a finite signed Borel measure. Denote further by $M_+(A)$ the cone of positive measures on A, by $FM(A)$ the space of finitely supported signed measures on A, and by $FM_+(A)$ the cone of finitely supported positive measures on A. The space $M(A)$ is equipped with the norm $\|\mu\| = |\mu|(A)$, where $|\mu| \in M_+(A)$ is the total variation of μ.

If (X, d) is a metric space and $A, B \subseteq X$, the **Hausdorff distance** between A and B is

$$d_H(A, B) = \max\Big\{\sup_{a\in A} d(a, B),\ \sup_{b\in B} d(b, A)\Big\};$$

it is easy to verify that this defines a metric on the class of compact subsets of X. If A and B are metric spaces then the **Gromov-Hausdorff distance** between A and B is

$$d_{GH}(A, B) = \inf d_H\big(\varphi(A), \psi(B)\big),$$

where the infimum is over all metric spaces X and isometric embeddings $\varphi\colon A \hookrightarrow X$ and $\psi\colon B \hookrightarrow X$. It is a nontrivial result that this defines a metric on the family of isometry classes of compact metric spaces; see [9, Chapter 3].

The precise normalizations used for Fourier transforms will not be important here, but for concreteness, the Fourier transform of a measure μ on $\mathbb{R}^n$ is defined as the function

$$\widehat{\mu}(\omega) = \int_{\mathbb{R}^n} e^{-i2\pi\langle x,\omega\rangle}\, d\mu(x),$$

where $\langle\cdot,\cdot\rangle$ is the standard inner product on $\mathbb{R}^n$, and the Fourier transform of $f \in L_1(\mathbb{R}^n)$ is the function

$$\widehat{f}(\omega) = \int_{\mathbb{R}^n} f(x) e^{-i2\pi\langle x,\omega\rangle}\, dx.$$

For $0 < p < \infty$, L_p will be used as shorthand for the vector space $L_p[0, 1]$ of equivalence classes (under almost everywhere equality) of measurable functions $x : [0, 1] \to \mathbb{R}$ such that

$$\|x\|_p = \left(\int_0^1 |x(t)|^p\ dt\right)^{1/p} < \infty.$$

As is well-known, $\|\cdot\|_p$ defines a quasinorm on L_p which is is only a norm when $p \geq 1$; it is less well-known that when $0 < p < 1$, $d(x, y) = \|x - y\|_p^p$ defines a metric on L_p. Below, L_p will be equipped with the metric $d(x, y) = \|x - y\|_p^{\min\{1,p\}}$ unless otherwise specified. An isometry involving these spaces is understood as a *metric*-preserving function and not a *quasinorm*-preserving function when $p < 1$ (see the comments following Proposition 3.4). Similarly, ℓ_p^n denotes $\mathbb{R}^n$ equipped with the the metric $d(x, y) = \|x - y\|_p^{\min\{1,p\}}$, where $\|x\|_p = \big(\sum_{j=1}^n |x_j|^p\big)^{1/p}$ for $0 < p < \infty$ and $\|x\|_\infty = \max_{1\leq j\leq n} |x_j|$.

Acknowledgements

I thank Tom Leinster for many interesting and enlightening conversations about the subject of this paper and related ideas. I also thank Carsten Schütt for pointing out the

results of [20], and the web sites the n-Category Café[1] and MathOverflow[2], which were central in starting my involvement in this project.

2. Magnitude of positive definite spaces

Three different definitions for the magnitude of an infinite metric space A have been proposed in [24, 25, 41, 42]. The first definition of $|A|$ is as the supremum of the magnitudes of finite subspaces of A (see [24, Definition 3.1.1]). This is unsatisfactory in general since magnitude is not monotone with respect to inclusion among arbitrary finite metric spaces (see [24, Example 2.1.7]), so that this definition is not consistent with the original one when restricted to finite spaces. However, [24, Corollary 2.4.4] shows that this is not the case among finite positive definite metric spaces, making this definition reasonable for compact positive definite metric spaces (the scope assumed in [24, Definition 3.1.1]).

The second approach, taken in [25] and [41], is to consider a sequence of finite subspaces $\{A_k\}$ such that $\lim_{k\to\infty} A_k = A$ in the Hausdorff distance, and then define $|A| = \lim_{k\to\infty} |A_k|$. This is unsatisfactory since it is not clear a priori whether this limit is independent of the approximating subspaces $\{A_k\}$.

The third approach, taken in [42], is to generalize the original definition of magnitude using measures for weightings. A **weight measure** for (A, d) is a finite signed measure $\mu \in M(A)$ such that $\int_A e^{-d(x,y)}\ d\mu(y) = 1$ for every $x \in A$. If A possesses a weight measure μ, then the magnitude of A may be defined as $\mu(A)$. If A possesses multiple weight measures, it is easy to check that they give the same value for the magnitude; however it is not clear how generally weight measures exist. (If A is a compact homogeneous space then A has a weight measure; see [42, Theorem 1]. Other sufficient conditions follow from Lemma 2.8 and Corollary 2.10 below.) In [42], Willerton showed that the magnitudes of intervals, circles, and Cantor sets, as computed via weight measures, coincide with their magnitudes as computed in [25] using the second approach.

The results of this section show that all these approaches to defining magnitude yield the same value of magnitude for compact positive definite metric spaces, and also develop some continuity properties of magnitude on such spaces. It will be convenient to take yet a fourth approach to the definition of magnitude, in terms of a Rayleigh-like quotient expression which already appears, in the finite case, in [24], and develop its relationships to the three approaches described above.

Given a compact metric space (A, d), define a bilinear form Z_A on $M(A)$ by

$$Z_A(\mu, \nu) = \int_A \int_A e^{-d(x,y)}\ d\mu(x)\ d\nu(y).$$

By Fubini's theorem, Z_A is symmetric. Observe that if μ is a weight measure for A and $\nu \in M(A)$, then $Z_A(\mu, \nu) = \nu(A)$.

If A is a compact positive definite metric space, the **magnitude** of is defined to be A to be

$$|A| = \sup \left\{ \frac{\mu(A)^2}{Z_A(\mu,\mu)} \middle| \mu \in M(A),\ Z_A(\mu,\mu) \neq 0 \right\}. \tag{2.1}$$

[1] `http://golem.ph.utexas.edu/category/`
[2] `http://mathoverflow.net`

In [24, Proposition 2.4.3] it is proved that when A is finite and positive definite, this coincides with the earlier definition (cf. Theorem 2.3 below). Observe that an immediate consequence of this definition is that if $B \subseteq A$ then $|B| \le |A|$.

It will also be useful to consider the quantity

$$|A|_+ = \sup \left\{ \frac{\mu(A)^2}{Z_A(\mu,\mu)} \middle| \mu \in M(A)_+,\ \mu \neq 0 \right\}. \tag{2.2}$$

Note that if A is positive definite, then $Z_A(\mu,\mu) > 0$ whenever μ is a nonzero positive measure. The quantity $|A|_+$ is called the **maximum diversity** of A because of an interpretation related to theoretical ecology (see [23] and the discussion at the end of [24, Section 2.4]). For any compact positive definite metric space A, it is easy to check that $|A|_+ \le \exp(\operatorname{diam}(A))$.

Compactness is a useful and natural-seeming condition to assume in this context. However, it is not clear that it is necessarily the most natural condition to use. If A is an infinite set in which each distinct pair of points is separated by a distance $r > 0$, then A is a *noncompact* positive definite metric space, which can nevertheless sensibly be assigned a finite magnitude e^r using the first definition proposed above (see [24, Section 3.1]). On the other hand, it is unknown at present whether the magnitude of a compact positive definite metric space can be infinite. As in [24, Section 3], attention will nevertheless be restricted here to compact spaces.

The following lemma is central to the results of this section. Recall that if (A,d) is a metric space and $f\colon A \to \mathbb{R}$ is uniformly continuous, the **modulus of continuity** of f is the function $\omega_f\colon (0,\infty) \to [0,\infty)$ defined by

$$\omega_f(\varepsilon) = \sup\big\{ |f(x) - f(y)| \,\big|\, x,y \in A,\ d(x,y) < \varepsilon \big\}.$$

Lemma 2.1. *Let A and B be compact subspaces of a metric space X and let $\mu \in M(A)$. For any $\varepsilon > d_H(A,B)$ there exists a $\nu \in M(B)$ such that $\nu(B) = \mu(A)$, $\|\nu\| \le \|\mu\|$, and for any uniformly continuous $f\colon X \to \mathbb{R}$,*

$$\left| \int_A f\ d\mu - \int_B f\ d\nu \right| \le \|\mu\|\, \omega_f(2\varepsilon).$$

Moreover, if μ is positive then ν can be taken to be positive.

Proof. Let $x_1, \dots, x_N \in A$ be the centers of open ε-balls which cover A. Each open ball $B(x_j,\varepsilon)$ in X contains a point $y_j \in B$, and then the balls $B(y_j, 2\varepsilon)$ cover A. Thus the disjoint open sets $U_1 = B(y_1, 2\varepsilon)$ and

$$U_j = B(y_j, 2\varepsilon) \setminus \bigcup_{k=1}^{j-1} B(y_k, 2\varepsilon) \quad \text{for } j = 2, \dots, N$$

also cover A. Let $\nu = \sum_{j=1}^N \mu(U_j \cap A)\delta_{y_j}$. Then

$$\begin{aligned}\left| \int_A f\ d\mu - \int_B f\ d\nu \right| &= \left| \sum_{j=1}^N \int_{U_j} \big(f(x) - f(y_j)\big)\ d\mu(x) \right| \le \sum_{j=1}^N \int_{U_j} |f(x) - f(y_j)|\ d\,|\mu|\,(x) \\ &\le \sum_{j=1}^N \omega_f(2\varepsilon)\,|\mu|\,(U_j \cap A) = \omega_f(2\varepsilon)\,|\mu|\,(A) = \|\mu\|\,\omega_f(2\varepsilon).\end{aligned}$$

Furthermore,

$$\nu(B) = \sum_{j=1}^{N} \mu(U_j \cap A) = \mu(A)$$

and

$$\|\nu\| = |\nu|(B) = \sum_{j=1}^{N} |\mu(U_j \cap A)| \leq \sum_{j=1}^{N} |\mu|(U_j \cap A) = |\mu|(A) = \|\mu\|. \qquad \square$$

Lemma 2.2. *A compact metric space A is positive semidefinite if and only if Z_A is a positive semidefinite bilinear form on $M(A)$. If Z_A is positive definite then A is positive definite.*

Proof. Recall that by definition (A, d) is positive (semi)definite if all of its finite subspaces are positive (semi)definite. The "if" parts follow by applying the positive (semi)definite bilinear form Z_A to finitely supported signed measures.

Now suppose that A is positive semidefinite and let $\mu \in M(A)$ and $\varepsilon > 0$. Apply Lemma 2.1 with B a finite $(\varepsilon/2)$-net in A to obtain $\nu \in M(B)$ with $\|\nu\| \leq \|\mu\|$ such that

$$\left| \int_A e^{-d(x,y)} \, d\mu(y) - \int_A e^{-d(x,y)} \, d\nu(y) \right| \leq \varepsilon \|\mu\|$$

for each $x \in A$. From this it follows that

$$\begin{aligned} |Z_A(\mu,\mu) - Z_A(\nu,\nu)| &\leq \left| \int_A \int_A e^{-d(x,y)} \, d\mu(y) \, d\mu(x) - \int_A \int_A e^{-d(x,y)} \, d\nu(y) \, d\mu(x) \right| \\ &\quad + \left| \int_A \int_A e^{-d(x,y)} \, d\mu(x) \, d\nu(y) - \int_A \int_A e^{-d(x,y)} \, d\nu(x) \, d\nu(y) \right| \\ &\leq \varepsilon \|\mu\|^2 + \varepsilon \|\mu\| \|\nu\| \leq 2\varepsilon \|\mu\|^2. \end{aligned}$$

Since B is a positive semidefinite finite metric space, $Z_A(\nu,\nu) \geq 0$, which implies $Z_A(\mu,\mu) \geq -2\varepsilon \|\mu\|^2$. Since $\varepsilon > 0$ was arbitrary, $Z_A(\mu,\mu) \geq 0$. $\square$

The next result, which generalizes [24, Proposition 2.4.3], shows the agreement of the present definition (2.1) with the measure-theoretic definition of magnitude used in [42], whenever both definitions can be applied.

Theorem 2.3. *Suppose A is a compact positive definite metric space. The supremum in* (2.1) *is achieved for a measure μ if and only if μ is a nonzero scalar multiple of a weight measure for A. If μ is a weight measure for A then $|A| = \mu(A)$.*

Proof. Suppose first that μ is a weight measure for A. (If μ achieves the supremum in (2.1), then so does any nonzero scalar multiple of μ by homogeneity.) By Lemma 2.2, Z_A is a positive semidefinite bilinear form on $M(A)$, and therefore satisfies the Cauchy-Schwarz inequality. Thus if $\nu \in M(A)$, then

$$\nu(A) = Z_A(\mu,\nu) \leq \sqrt{Z_A(\mu,\mu) Z_A(\nu,\nu)} = \sqrt{\mu(A) Z_A(\nu,\nu)},$$

with equality if $\nu = \mu$, and so

$$\mu(A) = \frac{\mu(A)^2}{Z_A(\mu,\mu)} = \sup \left\{ \frac{\nu(A)^2}{Z_A(\nu,\nu)} \middle| \nu \in M(A), \ Z_A(\nu,\nu) \neq 0 \right\}.$$

Now suppose that μ achieves the supremum in (2.1) and let $\nu \in M(A)$ satisfy $\nu(A) = 0$. Then for any $t \in \mathbb{R}$,

$$Z_A(\mu, \mu) \leq Z_A(\mu + t\nu, \mu + t\nu) = Z_A(\mu, \mu) + 2Z_A(\mu, \nu)t + Z_A(\nu, \nu)t^2.$$

Since $Z_A(\nu, \nu) \geq 0$ by Lemma 2.2, this implies that $Z_A(\mu, \nu) = 0$. Applying this in the case that $\nu = \delta_x - \delta_y$ for arbitrary $x, y \in A$ yields

$$\int_A e^{-d(x,z)} \, d\mu(z) = \int_A e^{-d(y,z)} \, d\mu(z),$$

and thus μ is a scalar multiple of a weight measure for A. □

Theorem 2.4 below shows that the present definition of magnitude (2.1) coincides, for positive definite spaces, with the first proposed definition discussed above.

Theorem 2.4. *For any positive definite compact metric space A,*

$$|A| = \sup \left\{ \frac{\mu(A)^2}{Z_A(\mu, \mu)} \middle| \mu \in FM(A), \ \mu \neq 0 \right\} = \sup\bigl\{|B| \bigm| B \subseteq A \text{ is finite}\bigr\}. \tag{2.3}$$

and

$$|A|_+ = \sup \left\{ \frac{\mu(A)^2}{Z_A(\mu, \mu)} \middle| \mu \in FM_+(A), \ \mu \neq 0 \right\} = \sup \left\{ |B|_+ \middle| B \subseteq A \text{ is finite} \right\}. \tag{2.4}$$

Proof. Observe first that when (A, d) is positive definite and $\mu \in FM(A)$, it follows that $Z_A(\mu, \mu) = 0$ only for $\mu = 0$. The second equality in (2.3) follows from [24, Proposition 2.4.3] (or Theorem 2.3 above), which shows that for finite positive definite spaces, the present definition (2.1) of magnitude agrees with the original definition. The second equality in (2.4) is immediate from (2.2).

In both (2.3) and (2.4) the first quantity is by definition greater than or equal to the second quantity. Let μ be a given measure on A such that $Z_A(\mu, \mu) \neq 0$, and let $\varepsilon > 0$. Apply Lemma 2.1 with B an $(\varepsilon/2)$-net in A to obtain $\nu \in FM(A)$, which is positive if μ is positive, such that $\nu(A) = \mu(A)$, $\|\nu\| \leq \|\mu\|$ and

$$\left| \int_A f \, d\mu - \int_A f \, d\nu \right| \leq \|\mu\| \, \omega_f(2\varepsilon)$$

for every continuous $f \colon A \to \mathbb{R}$.

Define $f_\mu \colon A \to \mathbb{R}$ by $f_\mu(x) = \int_A e^{-d(x,y)} \, d\mu(y)$, and define $f_\nu \colon A \to \mathbb{R}$ analogously. Then

$$|f_\nu(x) - f_\nu(y)| \leq \|\nu\| \, d(x, y),$$

and

$$|f_\mu(x) - f_\nu(x)| = \left| \int_A e^{-d(x,y)} \, d\mu(y) - \int_B e^{-d(x,y)} \, d\nu(y) \right| \leq 2 \|\mu\| \, \varepsilon.$$

Consequently,

$$\begin{aligned} |Z_A(\mu, \mu) - Z_B(\nu, \nu)| &= \left| \int_A f_\mu \, d\mu - \int_A f_\nu \, d\nu \right| \\ &\leq \left| \int_A f_\mu \, d\mu - \int_A f_\nu \, d\mu \right| + \left| \int_A f_\nu \, d\mu - \int_A f_\nu \, d\nu \right| \\ &\leq 2 \|\mu\|^2 \varepsilon + 2 \|\mu\| \, \|\nu\| \, \varepsilon \leq 4 \|\mu\|^2 \varepsilon. \end{aligned}$$

Therefore

$$\frac{\nu(A)^2}{Z_A(\nu,\nu)} = \frac{\mu(A)^2}{Z_A(\nu,\nu)} \geq \frac{\mu(A)^2}{Z_A(\mu,\mu) + 4\left\|\mu\right\|^2 \varepsilon}.$$

Since $\varepsilon > 0$ was arbitrary,

$$\sup\left\{\frac{\nu(A)^2}{Z_A(\nu,\nu)}\middle| \nu \in FM(A),\ \nu \neq 0\right\} \geq \frac{\mu(A)^2}{Z_A(\mu,\mu)},$$

and if μ is positive the same holds for the supremum over $FM_+(A)$. □

In some circumstances magnitude can be expressed in terms of functions instead of measures. A positive measure ρ on a metric space A is called a **good reference measure** if $\rho(U) > 0$ for every nonempty open $U \subseteq A$. For example, if $A \subseteq \mathbb{R}^n$ is the closure of its interior as a subset of $\mathbb{R}^n$, then Lebesgue measure restricted to A is a good reference measure. This may be generalized naturally in at least two ways. On the one hand, if A is a subset of a locally compact homogeneous metric space and A is the closure of its interior, then a Haar measure restricted to A is a good reference measure. On the other hand, if A is a metric space whose every nonempty open subset has the same Hausdorff dimension δ, then δ-dimensional Hausdorff measure is a good reference measure on A.

Given $h \in L_1(A,\rho)$, $h\rho$ denotes the signed measure on A defined by $(h\rho)(S) = \int_S h \, d\rho$. The proof of the following result is analogous to the proof of Theorem 2.4.

Proposition 2.5. *For any positive definite compact metric space A with a good reference measure ρ,*

$$|A| = \sup\left\{\frac{(h\rho)(A)^2}{Z_A(h\rho,h\rho)}\middle| h \in L_1(A,\rho),\ Z_A(h\rho,h\rho) \neq 0\right\}$$

and

$$|A|_+ = \sup\left\{\frac{(h\rho)(A)^2}{Z_A(h\rho,h\rho)}\middle| h \in L_1(A,\rho),\ h \geq 0,\ h \text{ is not } \rho\text{-a.e. } 0\right\}$$

The agreement of (2.1) with the definition of magnitude as the limit of magnitudes of an approximating sequence of subspaces, as in [25, 41], will follow from the next result, which is of independent interest.

Theorem 2.6. *The function $A \mapsto |A|$ (with values in $[1,\infty]$) is lower semicontinuous with respect to Gromov-Hausdorff distance on the class of compact positive definite metric spaces.*

Proof. Let (A,d) be a positive definite metric space with $|A| < \infty$ and let $\varepsilon > 0$ be given. (The case where $|A| = \infty$ is handled similarly.) Pick a signed measure μ on A with $Z_A(\mu,\mu) \neq 0$ such that

$$|A| \leq \frac{\mu(A)^2}{Z_A(\mu,\mu)}(1+\varepsilon).$$

Now let B be any other positive definite metric space with $d_{GH}(A,B) > 0$. Without loss of generality one may assume that $A, B \subseteq X$ for some metric space X, and $d_H(A,B) \leq 2d_{GH}(A,B)$. Let $\nu \in M(B)$ be as guaranteed by Lemma 2.1.

Define $f_\mu \colon X \to \mathbb{R}$ by $f_\mu(x) = \int_A e^{-d(x,y)} \, d\mu(y)$, and define f_ν analogously. Then

$$|f_\nu(x) - f_\nu(y)| \leq \|\nu\| \, d(x,y),$$

and by Lemma 2.1,

$$|f_\mu(x) - f_\nu(x)| = \left| \int_A e^{-d(x,y)} \, d\mu(y) - \int_B e^{-d(x,y)} \, d\nu(y) \right| \le 2 \|\mu\| \, d_H(A,B).$$

Consequently,

$$\begin{aligned} |Z_A(\mu,\mu) - Z_B(\nu,\nu)| &= \left| \int_A f_\mu \, d\mu - \int_B f_\nu \, d\nu \right| \\ &\le \left| \int_A f_\mu \, d\mu - \int_A f_\nu \, d\mu \right| + \left| \int_A f_\nu \, d\mu - \int_B f_\nu \, d\nu \right| \\ &\le 2 \|\mu\|^2 d_H(A,B) + 2 \|\mu\| \, \|\nu\| \, d_H(A,B) \\ &\le 8 \|\mu\|^2 d_{GH}(A,B). \end{aligned}$$

Therefore

$$\begin{aligned} |B| \ge \frac{\nu(B)^2}{Z_B(\nu,\nu)} &\ge \frac{\mu(A)^2}{Z_A(\mu,\mu) + 8 \|\mu\|^2 d_{GH}(A,B)} \\ &\ge \left(1 + \frac{8 \|\mu\|^2}{Z_A(\mu,\mu)} d_{GH}(A,B) \right)^{-1} \cdot \frac{|A|}{1+\varepsilon}. \end{aligned}$$

So if $d_{GH}(A,B) \le \frac{Z_A(\mu,\mu)}{8\|\mu\|^2}\varepsilon$, then $|B| \ge (1+\varepsilon)^{-2} |A|$. □

In general, $|A|$ is not a continuous function of A. Examples 2.2.8 and 2.4.9 of [24] give an example, due to S. Willerton, of a metric space A such that tA is positive definite for each $t > 0$ and $\lim_{t\to 0^+} |tA| = 6/5$, whereas $\lim_{t\to 0^+} tA = \{*\}$, which has magnitude 1. It is an open question whether $A \mapsto |A|$ is continuous when restricted to compact subsets of a fixed positive definite space. The *asymptotic conjectures* of [25] (see also [24, Conjecture 3.5.10]) would imply in particular that magnitude is continuous when restricted to compact convex subsets of ℓ_2^n.

Theorem 2.6 and the monotonicity of magnitude for positive definite spaces immediately imply the following result, which shows that the present definition (2.1) agrees with the definition of magnitude in terms of an approximating sequence of subspaces, and in particular shows that the latter definition is independent of the subspaces chosen.

Corollary 2.7. *If A is a compact positive definite metric space and $\{A_k\}$ is a sequence of compact subspaces of A such that* $\lim_{k\to\infty} d_H(A_k, A) = 0$, *then* $|A| = \lim_{k\to\infty} |A_k|$.

Theorems 2.3 and 2.4 and Corollary 2.7 completely explain the agreement of calculations of magnitudes using different definitions in [25] and [42], since all of the spaces involved are positive definite by [24, Proposition 2.4.13] and Theorem 3.6 below.

The remaining results of this section develop some additional properties of maximum diversity, which yield information about magnitude for a particular class of spaces. A compact positive definite metric space A is called **positively weighted** if $|A| = |A|_+$. A finite positive definite space is positively weighted if and only if its weighting has only nonnegative components; several results in [24, Section 2.4] give sufficient conditions for this property and properties of the magnitude of such finite spaces.

The following lemma gives several sufficient conditions for a compact positive definite metric space to be positively weighted. It will be seen in Corollary 2.10 that the first of these conditions is also necessary.

Lemma 2.8. *Let (A, d) be a compact positive definite metric space. Under any of the following conditions, A is positively weighted.*

1. *There is a positive weight measure on A.*
2. *The isometry group of A acts transitively on the points of A (i.e., A is a homogeneous space).*
3. *Every finite subset of A has a weighting with only nonnegative components.*
4. *There is an isometric embedding of A into $\mathbb{R}$, with the standard metric on $\mathbb{R}$.*
5. *For every $x, y, z \in A$, $d(x, y) \leq \max\{d(x, z), d(z, y)\}$ (i.e., A is an ultrametric space).*

Proof. (1) This follows from Theorem 2.3.

(2) A compact homogeneous space has a nonnegative weight measure (see [42, Theorem 1] and the comments following it).

(3) This follows from Theorem 2.4.

(4) By [25, Theorem 4] or [24, Proposition 2.4.13], every finite subset of $\mathbb{R}$ has a nonnegative weighting.

(5) By [24, Proposition 2.4.18] (originally proved in [30]), every finite ultrametric space has a nonnegative weighting. □

It should be noted that every subset of $\mathbb{R}$ and every ultrametric space is positive definite (see [24, Proposition 2.4.13] and [36], respectively; also [24, Proposition 2.4.18] and Theorem 3.6 below). There exist homogeneous spaces which are not positive definite (see [24, Example 2.1.7]). Example 2.4.16 in [24] shows that when $n \geq 2$, not all compact subsets of ℓ_1^n are positively weighted, and numerical calculations in [41] show that not all compact subsets of ℓ_2^n are positively weighted, although ℓ_1^n and ℓ_2^n are positive definite (see [24, Theorems 2.4.14 and 2.5.3]; also Theorem 3.6 below).

Proposition 2.9. *If A is a compact positive definite metric space, then the supremum in the definition (2.2) of $|A|_+$ is achieved by some $\mu \in M_+(A)$.*

Proof. Denote by $P(A) = \{\mu \in M_+(A) \mid \mu(A) = 1\}$ the space of probability measures on A. A well-known consequence of the Banach-Alaoglu theorem is that when A is compact, $P(A)$ is compact with respect to the weak-$*$ topology inherited from the action of $M(A)$ as the dual of the Banach space $(C(A), \|\cdot\|)$. This topology on $P(A)$ is metrized by the Wasserstein distance

$$d_W(\mu, \nu) = \sup \left\{ \int_A f \, d\mu - \int_A f \, d\nu \,\middle|\, f \colon A \to \mathbb{R} \text{ is 1-Lipschitz} \right\}$$

(see e.g. [37, Corollary 6.12]). By homogeneity,

$$|A|_+ = \sup_{\mu \in P(A)} \frac{1}{Z_A(\mu, \mu)}. \tag{2.5}$$

In the notation of the proof of Theorem 2.6, for $\mu, \nu \in P(A)$,

$$|Z_A(\mu, \mu) - Z_A(\nu, \nu)| \leq \left| \int_A f_\mu \, d\mu - \int_A f_\mu \, d\nu \right| + \left| \int_A f_\nu \, d\mu - \int_A f_\nu \, d\nu \right|$$

since $\int_A f_\mu \, d\nu = \int_A f_\nu \, d\mu$ by Fubini's theorem. As seen in the proof of Theorem 2.6, f_μ and f_ν are both 1-Lipschitz (because $\|\mu\| = \|\nu\| = 1$), and therefore

$$|Z_A(\mu,\mu) - Z_A(\nu,\nu)| \leq 2 d_W(\mu,\nu).$$

By continuity and compactness, the supremum in (2.5) is achieved by some $\mu \in P(A)$, which then also achieves the supremum in (2.2). □

Corollary 2.10. *If A is a positively weighted compact positive definite metric space, then there is a positive weight measure on A.*

Proof. By Proposition 2.9, there is a $\mu \in M_+(A)$ which achieves the supremum in (2.1). By Theorem 2.3, μ is, up to a scalar multiple, a weight measure for A. □

Proposition 2.11. *The function $A \mapsto |A|_+$ is continuous with respect to the Gromov-Hausdorff distance on the class of compact positive definite metric spaces.*

Proof. Lower semicontinuity follows exactly as in the proof of Theorem 2.6. The argument for upper semicontinuity proceeds along similar lines.

Given A and B with $d_{GH}(A,B) > 0$, we may assume that $A, B \subseteq X$ for some metric space X and $0 < d_H(A,B) \leq 2 d_{GH}(A,B)$. Suppose that $\mu \in M_+(B)$ satisfies

$$|B|_+ = \frac{\mu(B)}{Z_B(\mu,\mu)}$$

(as guaranteed by Proposition 2.9) and construct $\nu \in M_+(A)$ as in Lemma 2.1. Proceeding analogously to the proof of Theorem 2.6, one obtains that

$$\frac{\mu(B)^2}{Z_B(\mu,\mu)} \leq \frac{\nu(A)^2}{Z_A(\nu,\nu) - 8\|\mu\|^2 d_{GH}(A,B)}.$$

Since μ is positive, $\|\mu\| = \mu(B) = \nu(A)$, and so

$$\begin{aligned} |B|_+ &\leq \frac{\nu(A)^2}{Z_A(\nu,\nu) - 8\nu(A)^2 d_{GH}(A,B)} \leq \left(1 - \frac{8\nu(A)^2}{Z_A(\nu,\nu)} d_{GH}(A,B)\right)^{-1} |A|_+ \\ &\leq \left(1 - 8|A|_+ d_{GH}(A,B)\right)^{-1} |A|_+ . \end{aligned}$$

So if $d_{GH}(A,B) \leq \frac{\varepsilon}{8|A|_+}$ for $0 < \varepsilon < 1$, then $|B|_+ \leq (1-\varepsilon)^{-1} |A|_+$. □

Corollary 2.12. *Magnitude is continuous with respect to the Gromov-Hausdorff distance on the class of positively weighted compact positive definite metric spaces.*

A particular case of Corollary 2.12 is that magnitude is continuous with respect to Gromov-Hausdorff distance for compact subsets of $\mathbb{R}$. In this setting, the slightly weaker result of continuity with respect to Hausdorff distance also follows from an exact integral formula for the magnitude of a compact subset of $\mathbb{R}$ given in [24, Proposition 3.2.3].

3. Examples and counterexamples of positive definite spaces

This section is divided into two subsections. The first investigates sufficient conditions for positive definiteness of a metric space, in particular relating it to the classical property of negative type, and gives a number of examples of positive definite metric spaces. The second subsection gives some examples of metric spaces which are not positive definite, in particular demonstrating that some natural operations on metric spaces do not preserve positive definiteness.

3.1. Sufficient conditions for positive definiteness.

A function $f\colon E \to \mathbb{C}$ on a vector space E is called **positive definite** if, for every finite nonempty $A \subseteq E$, the matrix $[f(x-y)]_{x,y\in A} \in \mathbb{C}^{A\times A}$ is positive *semidefinite*; f is called **strictly positive definite** if $[f(x-y)]_{x,y\in A}$ is positive definite. (This inconsistency in terminology is unfortunately well-established.) Thus a translation-invariant metric d on E is a positive definite (respectively, positive semidefinite) metric if and only if $x \mapsto e^{-d(x,0)}$ is a strictly positive definite (positive definite) function.

The following classical result connects positive definiteness to harmonic analysis (see e.g. [14]).

Proposition 3.1 (Bochner's theorem)**.** *A continuous function $f\colon \mathbb{R}^n \to \mathbb{C}$ is positive definite if and only if $f = \widehat{\mu}$ for some positive measure μ on $\mathbb{R}^n$.*

Bochner's theorem does not consider *strictly* positive definite functions, and thus cannot directly identify positive definite metrics on $\mathbb{R}^n$. A theory of strictly positive definite functions is presented in [40, Chapter 6], which contains several counterparts to Bochner's theorem, including the following.

Proposition 3.2 ([40, Theorem 6.11])**.** *Suppose $f \in L_1(\mathbb{R}^n)$ is continuous. Then f is strictly positive definite if and only if f is bounded and $\widehat{f}$ is nonnegative and not uniformly 0.*

Rather than being applied directly here, Proposition 3.2 will be combined with classical results to prove Theorem 3.3 below, which allows positive definiteness of a metric space A to be deduced from positive *semidefiniteness* of rescalings of A.

A metric space A is **stably positive (semi)definite** if tA is positive (semi)definite for every $t > 0$. Since the definition of magnitude implicitly involves an arbitrary choice of scale (as discussed at the beginning of [24, Section 2.2]), stable positive definiteness is arguably a more natural condition on a metric space than positive definiteness. The space A is of **negative type** if $A^{1/2}$ is isometric to a subset of a Hilbert space. Spaces of negative type have been studied extensively in the theory of embeddings of metric spaces; see e.g. [39, 3], or [5, Section 2] for a concise recent survey. (The terminology *negative type* stems from an alternative characterization of such spaces which will not be needed here.)

The following result shows that the theory of magnitude leads naturally to the classical notion of negative type, and that the literature on negative type gives many examples of positive definite metric spaces. As mentioned above, it is also a useful tool for upgrading positive semidefiniteness to positive definiteness, allowing one to avoid using Proposition 3.2 or related results explicitly.

Theorem 3.3. *The following are equivalent for a metric space A.*

1. *A is stably positive definite, and thus every compact subset of A has a defined (possibly infinite) magnitude.*

(2) *A is stably positive semidefinite.*
(3) *There is a sequence* $\{t_k > 0 \mid k \in \mathbb{N}\}$ *with* $\lim_{k\to\infty} t_k = 0$ *such that* $t_k A$ *is positive semidefinite for every* k.
(4) *A is of negative type.*

Proof. The implications (1) ⇒ (2) ⇒ (3) are trivial. The equivalence of (2) and (4) was proved in [33, Theorem 1] (although that paper was written long before the terminology used here was introduced). The proof will be completed by showing that (3) ⇒ (2) and (4) ⇒ (1).

Assume the condition in (3). It suffices to assume that A is finite. A finite metric space A is positive semidefinite if and only if $\lambda_{\min}(\zeta_A) \geq 0$, where $\lambda_{\min}$ denotes the smallest eigenvalue of a symmetric matrix. Since the smallest eigenvalue of a symmetric matrix is a continuous function of the matrix entries (see e.g. [12, p. 370]), the set of t such that tA is positive semidefinite is a closed subset of $(0, \infty)$.

If A is a finite positive semidefinite metric space, then by the Schur product theorem [12, Theorem 7.5.3], ζ_{nA} is a positive semidefinite matrix for every positive integer n. Therefore the condition in (3) implies that the set of t such that tA is positive semidefinite is a dense subset of $(0, \infty)$, and so A is stably positive semidefinite.

Suppose finally that (A, d) is of negative type. It suffices again to assume that A is finite. Then there is a function $\varphi\colon A \to \ell_2^n$ for some n such that $d(x, y) = \|\varphi(x) - \varphi(y)\|_2^2$ for every $x, y \in A$, so in particular φ is injective. For $t > 0$ define $f\colon \mathbb{R}^n \to \mathbb{R}$ by $f(x) = \exp\bigl(-t \|x\|_2^2\bigr)$. Proposition 3.2 implies that f is a strictly positive definite function, which means that

$$\zeta_{tA} = \bigl[\exp(-td(x, y))\bigr]_{x,y\in A} = \bigl[f\bigl(\varphi(x) - \varphi(y)\bigr)\bigr]_{x,y\in A}$$

is a positive definite matrix. Thus A is stably positive definite. □

The equivalence of negative type with stable positive *semidefiniteness*, due to Schoenberg [33], is well-known to experts on embeddings of metric spaces, and has been generalized in various directions. The theory of magnitude, however, requires positive *definiteness*, and the equivalence of stable positive definiteness with negative type appears to be new. A further equivalence, between negative type and Enflo's notion of generalized roundness [6], was proved in [27].

After Theorem 3.3 was first proved (with a slightly weaker version of condition (3)), T. Leinster found a direct proof of the implication (3) ⇒ (1). This gives an alternative, more elementary proof of the equivalence of conditions (1)–(3) which is independent of the notion of negative type. Leinster's argument furthermore obviates the need for Proposition 3.2, or indeed any mention of *strictly* positive definite functions, in the development of this theory. (However, the results of Section 4 below require properties of certain functions which are stronger than strict positive definiteness.) Leinster's proof is included here with his permission.

Second proof of Theorem 3.3, (3) ⇒ (1). Assume without loss of generality that A is finite. By [24, Proposition 2.2.6 (i)], ζ_{tA} is invertible for all but finitely many values of $t > 0$, and thus there is an $\varepsilon > 0$ such that ζ_{tA} is invertible for all $t \in (0, \varepsilon)$. By the continuity property used above, either (a) $\lambda_{\min}(\zeta_{tA}) < 0$ for all $t \in (0, \varepsilon)$ or (b) $\lambda_{\min}(\zeta_{tA}) > 0$ for all $t \in (0, \varepsilon)$. By condition (3), (a) is impossible, so (b) holds. Thus tA is positive definite for all $t \in (0, \varepsilon)$. By the positive *definite* version of the Schur product theorem, ntA is positive definite for every $t \in (0, \varepsilon)$ and positive integer n, and thus A is stably positive definite. □

The remainder of this subsection is devoted to collecting examples of metric spaces which are of negative type and hence positive definite. The following result essentially goes back to Lévy, generalized by Bretagnolle, Dacunha-Castelle, and Krivine [2]. For the precise version stated here, see [1, Theorem 8.9].

Proposition 3.4. *Let $(E, \|\cdot\|_E)$ be a real separable quasinormed space and let $0 < p \leq 2$. The following are equivalent.*

1. *The function $f(x) = \exp(-\|x\|_E^p)$ is positive definite on E.*
2. *There is a linear map $T\colon E \to L_p$ such that $\|x\|_E = \|T(x)\|_p$ for every $x \in E$.*

Some remarks are in order at this point. First, a map as in part (2) of Proposition 3.4 is usually called an isometry between quasinormed spaces, even when the quasinorms are not norms. The slightly pedantic formulation of Proposition 3.4 above is to avoid ambiguous uses of the word isometry. Observe that if a quasinormed space $(E, \|\cdot\|_E)$ satisfies the conditions of Proposition 3.4 for some $p \in (0, 1]$, then $d(x, y) = \|x - y\|_E^p$ defines a metric on E and the map T is an isometry from E equipped with this metric into L_p with the metric $d(x, y) = \|x - y\|_p^p$.

Second, the restriction to separable spaces here is merely for convenience of exposition (in order to avoid introducing nonseparable L_p spaces). The main interest here is whether compact subsets of E are positive definite, so this is no real restriction.

Finally, if a vector space E has a metric d which is homogeneous of any degree (in the sense that for some $\beta > 0$, $d(tx, ty) = t^\beta d(x, y)$ for every $x, y \in E$ and $t > 0$), then $tE = (E, td)$ is isometric to E for every $t > 0$. It follows from Proposition 3.3 that (E, d) is stably positive definite if and only if (E, d) is positive semidefinite.

With these remarks in mind, the following is an immediate consequence of Proposition 3.4 and Theorem 3.3.

Corollary 3.5. *The following are equivalent for a real separable normed space $(E, \|\cdot\|_E)$.*

1. *E is a positive semidefinite metric space.*
2. *E is a positive definite metric space.*
3. *E is a metric space of negative type.*
4. *E is isometric to a linear subspace of L_1.*

The next result collects several large classes of metric spaces which are known to have negative type, and whose compact subspaces therefore all have well-defined magnitudes by Theorem 3.3. The list is not intended to be exhaustive.

Theorem 3.6. *Any metric space from each of the following classes is of negative type, and hence positive definite.*

1. *A^α, where A is a metric space of negative type and $0 < \alpha \leq 1$.*
2. *L_p for $0 < p \leq 2$ (with the metric $d(x, y) = \|x - y\|_p^{\min\{1,p\}}$).*
3. *Two-dimensional real normed spaces.*
4. *Metric spaces with at most four points.*
5. *Ultrametric spaces.*
6. *Round spheres (with the geodesic distance).*
7. *Real or complex hyperbolic space.*
8. *Weighted trees.*

Proof. (1) This was proved by Schoenberg [33, p. 527].

(2) It was proved in [2] that if $1 \le p \le 2$, then L_p is isometric to a subspace of L_1. Thus it suffices to prove the claim in the case $0 < p \le 1$.

Suppose now that $0 < p \le 1$. The function $f(x) = \exp(-\|x\|_p^p) = e^{-d(x,0)}$ is positive definite on L_p by Proposition 3.4, which as observed above implies that L_p is positive semidefinite and hence (by homogeneity and Theorem 3.3) of negative type.

(3) By [19, Corollary 6.8], every two-dimensional real normed space is isometric to a subspace of L_1.

(4) By [43, Theorem 1], every N-point metric space with $N \ge 4$ embeds isometrically into ℓ_∞^{N-2}. In particular, every four-point space is isometric to a subset of ℓ_∞^2, which is two-dimensional so part (3) applies, or more simply is isometric to ℓ_1^2.

(5) Every finite ultrametric space embeds isometrically in ℓ_2^n for some n; see e.g. [26].

(6) This follows from results in [15, 16]; see [11, p. 263].

(7) This is proved in [7, Corollaires 7.4 and 7.7].

(8) This is proved in [11, Corollary 7.2]. □

The class of functions of metrics which preserve negative type, as $t \mapsto t^\alpha$ does in part (1) above, was determined in [32].

Leinster proved directly that ℓ_1^n and ℓ_2^n are positive definite for every n in Propositions 2.4.14 and 2.5.3 of [24], respectively. Leinster's proof for ℓ_2^n is based on the same idea behind the proof of Proposition 3.2 which partly underlies the (first) proof of Theorem 3.3.

The space L_0 of measurable functions $f\colon [0,1] \to \mathbb{R}$ is also of negative type when equipped with an appropriate metric that metrizes the topology of convergence in measure; see [1, p. 187].

A direct proof that three-point metric spaces are positive definite is given in [24, Proposition 2.4.15]. No proof that every four-point space has a defined magnitude is known which does not rely on an embedding into a positive definite normed space.

Ultrametric spaces were directly proved to be positive definite in [36], see also [24, Proposition 2.4.18].

Some necessary conditions for manifolds to have negative type are also known. For example, a compact Riemannian manifold of dimension at least two of negative type must be simply connected [10, Theorem 5.4], and a compact Riemannian symmetric space of negative type must be a round sphere [17, Corollary 2.6]. Since compact symmetric spaces are homogeneous, their magnitude can nevertheless be defined via weight measures as in [42].

3.2. **Counterexamples.**

This subsection collects several examples which demonstrate the limits of some of the results in the previous two subsections, or show that some appealing conjectures about positive definiteness are false. Many of these and related examples are known in the literature on metric spaces of negative type.

The first example shows that the converse of Theorem 3.6 (1) is false. Koldobsky [18] constructed a normed space E which embeds as a quasinormed space (i.e. in the sense of Proposition 3.4 (2)) into $L_{1/2}$ but does not embed in L_1. By Proposition 3.4 and homogeneity, this means that the metric space $E^{1/2}$ is stably positive definite, but E is not positive semidefinite.

The second example shows that the threshold $p = 2$ in Theorem 3.6 (2) and the dimension two in Theorem 3.6 (3) are both optimal. It was proved by Dor [4] that if $2 < p \le \infty$ then ℓ_p^3 does not embed isometrically in L_1, and is therefore not positive definite. In particular, L_p is not positive definite for any $p > 2$.

The third example shows that the cardinality four in Theorem 3.6 (4) is optimal. As shown in [24, Example 2.2.7], the vertices of the complete bipartite graph $K_{3,2}$, with the shortest path metric, form a metric space which is not positive definite if all edges have equal lengths $r < \log \sqrt{2}$.

The fourth set of examples are compact Riemannian manifolds which are not positive definite with the geodesic metric. By [10, Theorem 5.4], if M is any non-simply connected compact Riemannian manifold M of dimension at least two, then M fails to have negative type. Thus by Theorem 3.3 there is some $t > 0$ such that tM is not positive definite.

It is also possible to give an example which is both more elementary is topologically a sphere. The idea is simply to construct a surface S which almost isometrically contains a copy of $K_{3,2}$ with short edge lengths.

Start with a 2-sphere with radius smaller than $(\log 2)/\pi$ in $\mathbb{R}^3$, and consider the following five points: two opposite poles, and three equidistant points on the equator. Drawing the lines of longitude through the latter three points, we obtain a copy of $K_{3,2}$ in the sphere with equal edge lengths $r < \log \sqrt{2}$. Now put large bulges on the sphere in the three regions delineated by the three lines of longitude. These bulges may be made large enough to make the geodesic distance on S between the three equatorial points arbitrarily close to the graph distance $2r$.

Since $K_{3,2}$ with edge lengths r is not positive semidefinite, and the smallest eigenvalue of a symmetric matrix is a continuous function of the matrix entries, it follows that the bilinear form Z_S on $M(S)$ is not positive semidefinite.

The class of positive definite metric spaces is closed under taking ℓ_1 products [24, Lemma 2.4.2 (ii)]. The next set of examples shows that ℓ_p products for any $p > 1$ fail to preserve positive definiteness, even in the more restricted context of positive definite normed spaces (in which one usually speaks of ℓ_p *sums*). For $p > 2$, this follows from the fact discussed above that ℓ_p^3, which is the ℓ_p product of three copies of the positive definite space $\mathbb{R}$, is not positive definite. For $1 < p \le 2$, [20, Corollary 3.4] shows that the ℓ_p sum of n copies of ℓ_1^n does not embed isometrically in L_1 when n is sufficiently large (specifically, when $n > (5\sqrt{2})^{p/(p-1)}$; this bound is not sharp), and hence is not positive semidefinite.

For a small concrete example, let $A = \{0, \pm e_1, \pm e_2\} \subseteq \ell_1^2$, and consider $A \times A \subseteq \ell_1^2 \oplus_2 \ell_1^2$.. Then numerical calculation of the eigenvalues of $\zeta_{t(A\times A)}$ for small t shows that $A \times A$ is not stably positive definite.

The last examples show that the property of positive definiteness has no simple relationship with 1-Lipschitz maps. First, positive definiteness is not preserved by 1-Lipschitz maps. In fact, since every separable Banach space is isometric to a quotient space of the positive definite space ℓ_1 (cf. [28, p. 108]), it happens quite generically that 1-Lipschitz maps fail to preserve positive definiteness. For a concrete low-dimensional example, one can define a linear such map $\ell_1^4 \to \ell_\infty^3$ (recall that the former space is positive definite and the latter is not) by mapping the standard basis vectors of ℓ_1^4 to the four vertices on one facet of the ℓ_∞^3 unit ball.

On the other hand, if B is a positive definite metric space and there is a 1-Lipschitz surjection $f\colon A \to B$, A need not be positive definite. In the language of [24, Definition 2.2.4], this means that positive definiteness is not preserved by expansions. One can even insist that f be a bijection. Let E be a normed space which is not positive definite and let $A \subseteq E$ be any finite subset which is not positive definite. Then a generic linear functional $f\colon E \to \mathbb{R}$ is injective when restricted to A, and $B = f(A) \subseteq \mathbb{R}$ is positive definite.

4. Magnitude in ℓ_p^n

This section generalizes some results of Leinster [24] about the magnitude of subsets of ℓ_1^n and ℓ_2^n to the spaces $\left(\ell_p^n\right)^\alpha$ for $0 < p \le 2$ and $0 < \alpha \le 1$. Leinster's proofs for Euclidean space are based on the exact formula (4.4) for the Fourier transform of the function $x \mapsto e^{-\|x\|_2}$ on $\mathbb{R}^n$ (for ℓ_1^n more elementary tools suffice). For $0 < p \le 2$, $0 < \alpha \le 1$, and a positive integer n, define

$$F_{p,\alpha}^n\colon \mathbb{R}^n \to \mathbb{R}, \qquad F_{p,\alpha}^n(x) = \exp\bigl(- \|x\|_p^{\alpha \min\{1,p\}}\bigr).$$

The generalizations here require proving that the Fourier transforms $\widehat{F_{p,\alpha}^n}$ share the properties of $\widehat{F_{2,1}^n}$ which are essential in Leinster's proofs.

Parts (1) and (2) of Theorem 3.6 imply that $\left(\ell_p^n\right)^\alpha$ is of negative type, which by Theorem 3.3 is equivalent to the statement that $F_{p,\alpha}^n$ is a strictly positive definite function. Lemma 4.1 is a quantitative sharpening of this fact. In probabilistic terms, it gives polynomial lower bounds on the densities of a particular class of stable random vectors. As such, it may already be known in the probability literature, although we have been unable to find a statement. The bounds given by the proof are nonoptimal, but are sufficient for the purposes of the present paper.

Lemma 4.1. *Given $0 < p \le 2$, $0 < \alpha \le 1$, and a positive integer n there is a constant $c_{p,\alpha,n} > 0$ such that*

$$\widehat{F_{p,\alpha}^n}(\omega) \ge c_{p,\alpha,n}\bigl(1 + \|\omega\|_2\bigr)^{-(1+p)n}$$

for every $\omega \in \mathbb{R}^n$.

Proof. Suppose first that $0 < p < 2$. Define $\gamma_p\colon \mathbb{R} \to \mathbb{R}$ by $\gamma_p(t) = e^{-|t|^p}$. Then [19, Lemma 2.27] shows that $\widehat{\gamma_p} > 0$ everywhere, and asymptotic expansions in Theorems 2.4.1, 2.4.2, and 2.4.3 in [13] imply that

$$\lim_{|\omega|\to\infty} \widehat{\gamma_p}(\omega)\, |\omega|^{1+p}$$

exists and is finite. It follows that there is a constant $c(p) > 0$ such that

$$\widehat{\gamma_p}(\omega) \ge c(p)(1 + |\omega|)^{-(1+p)} \tag{4.1}$$

for every $\omega \in \mathbb{R}$.

A theorem of Bernstein [8, Theorem XIII.4] implies that for every $r \in (0, 1]$ there is a probability measure μ_r on $[0, \infty)$ such that

$$e^{-t^r} = \int_0^\infty e^{-ts}\ d\mu_r(s) \tag{4.2}$$

for every $t \geq 0$. In particular, for $r = \alpha \min\{1/p, 1\}$,

$$F^n_{p,\alpha}(x) = \int_0^\infty e^{-s\|x\|_p^p} \, d\mu_r(s) = \int_0^\infty \left(\prod_{j=1}^n \gamma_p\big(s^{1/p} x_j\big) \right) \, d\mu_r(s).$$

Now by Fubini's theorem, a linear change of variables, and (4.1),

$$\begin{aligned} \widehat{F^n_{p,\alpha}}(\omega) &= \int_0^\infty \left(\prod_{j=1}^n s^{-1/p} \widehat{\gamma_p}\big(s^{-1/p}\omega_j\big) \right) \, d\mu_r(s) \\ &\geq c(p)^n \int_0^\infty \left(s^{-1/p}\big(1 + s^{-1/p}\left\|\omega\right\|_2\big)^{-(1+p)} \right)^n \, d\mu_r(s). \end{aligned} \tag{4.3}$$

If $0 \leq s \leq 1$, then

$$s^{-1/p}\big(1 + s^{-1/p}\left\|\omega\right\|_2\big)^{-(1+p)} = s\big(s^{1/p} + \left\|\omega\right\|_2\big)^{-(1+p)} \geq s\big(1 + \left\|\omega\right\|_2\big)^{-(1+p)}.$$

If $s > 1$, then

$$s^{-1/p}\big(1 + s^{-1/p}\left\|\omega\right\|_2\big)^{-(1+p)} \geq s^{-1/p}\big(1 + \left\|\omega\right\|_2\big)^{-(1+p)}.$$

Thus

$$\widehat{F^n_{p,q}(\omega)} \geq c(p)^n \left(\int_{[0,1]} s^n \, d\mu_r(s) + \int_{(1,\infty)} s^{-n/p} \, d\mu_r(s) \right) \big(1 + \left\|\omega\right\|_2\big)^{-(1+p)d}.$$

Now suppose $p = 2$. When also $\alpha = 1$, there is the exact formula

$$\widehat{F^n_{2,1}}(\omega) = \frac{c_n}{\big(c_n' + \left\|\omega\right\|_2^2\big)^{(n+1)/2}}, \tag{4.4}$$

where $c_n, c_n' > 0$ are constants (which can be given explicitly) depending only on n (see [35, Theorem I.1.4]). By (4.2), Fubini's theorem, a linear change of variables, and (4.4),

$$\widehat{F^n_{2,\alpha}}(\omega) = \int_0^\infty s^{-n} \frac{c_n}{\big(c_n' + s^{-2}\left\|\omega\right\|_2^2\big)^{(n+1)/2}} \, d\mu_\alpha(s). \tag{4.5}$$

The proof is completed as before. □

Lemma 4.2. *Given $0 < p \leq 2$, $0 < \alpha \leq 1$, and a positive integer n, for each $\omega \in \mathbb{R}^n$, the function $t \mapsto \widehat{F^n_{p,\alpha}}(t\omega)$ is decreasing for $t \geq 0$.*

Proof. For $p = 2$ the lemma follows from (4.5) in the proof of Lemma 4.1.

For $p < 2$, [13, Theorem 2.5.3] implies that for any $\omega \in \mathbb{R}$, $t \mapsto \widehat{\gamma_p}(t\omega)$ is decreasing on $[0, \infty)$. The lemma in this case follows from this fact and the equality in (4.3). □

The main results of this section, Theorems 4.3 and 4.4, were proved by Leinster in the cases that $\alpha = 1$ and $p = 1, 2$ (see Theorem 3.4.8, Proposition 3.5.3, and Theorem 3.5.5 in [24]). The proofs below generalize Leinster's proofs for ℓ_2^n, using Lemmas 4.1 and 4.2 in place of the exact formula (4.4). The results of Section 2 allow the exposition to be simplified somewhat by working with measures instead of finite subsets.

Theorem 4.3. *Let A be a compact subset of $\big(\ell_p^n\big)^\alpha$, where $0 < p \leq 2$ and $0 < \alpha \leq 1$. Then $|A| < \infty$.*

Proof. Let $\psi\colon \mathbb{R}^n \to \mathbb{R}$ be an even, compactly supported, C^∞ function such that $\psi(x) = 1$ for all $x \in A - A = \{y - z \mid y, z \in A\}$. Then $\widehat{\psi}$ is a real-valued Schwartz function. By Lemma 4.1 there is some constant $C(p, \alpha, \psi) > 0$ such that

$$\widehat{\psi} \le C(p, \alpha, \psi)\widehat{F^n_{p,\alpha}}$$

everywhere on $\mathbb{R}^n$.

Since $F^n_{p,\alpha}$ is positive definite, integrable, and continuous at 0, [35, Corollary I.1.26] implies that $\widehat{F^n_{p,\alpha}} \in L_1(\mathbb{R}^n)$ and thus (since $\widehat{F^n_{p,\alpha}}$ is also even) that $F^n_{p,\alpha}$ is the Fourier transform of $\widehat{F^n_{p,\alpha}}$. Now for any $\mu \in M(A)$, by Fubini's theorem,

$$\begin{aligned}
Z_A(\mu, \mu) &= \int_A \int_A F^n_{p,\alpha}(x - y)\ d\mu(x)\ d\mu(y) \\
&= \int_A \int_A \int_{\mathbb{R}^n} \widehat{F^n_{p,\alpha}}(\omega) e^{-i2\pi\langle x-y,\omega\rangle}\ d\omega\ d\mu(x)\ d\mu(y) \\
&= \int_{\mathbb{R}^n} |\widehat{\mu}(\omega)|^2\, \widehat{F^n_{p,\alpha}}(\omega)\ d\omega \\
&\ge \frac{1}{C(p, \alpha, \psi)} \int_{\mathbb{R}^n} |\widehat{\mu}(\omega)|^2\, \widehat{\psi}(\omega)\ d\omega \\
&= \frac{1}{C(p, \alpha, \psi)} \int_A \int_A \psi(x - y)\ d\mu(x)\ d\mu(y) = \frac{1}{C(p, \alpha, \psi)} \mu(A)^2.
\end{aligned}$$

Thus by (2.1), $|A| \le C(p, \alpha, \psi)$. □

It is natural, especially in light of Theorem 3.6, to ask whether Theorem 4.3 applies to $(E)^\alpha$ for arbitrary finite dimensional subspaces E of L_p. To extend the present proof to this setting would require generalizing a lower bound as in Lemma 4.1 to the densities of much more general stable random vectors. Although such bounds have been the subject of much study (see e.g. [38]) it appears that known results do not suffice for this purpose.

Theorem 4.4. *Let A be a compact subset of $\left(\ell_p^n\right)^\alpha$, where $0 < p \le 2$ and $0 < \alpha \le 1$. There is a constant $C > 0$ such that $|tA| \le Ct^n$ for all $t \ge 1$.*

Proof. Let ψ be as in the proof of Theorem 4.3, and for $t \ge 1$ define $\psi_t(x) = \psi(\frac{x}{t})$, so that ψ_t is an even, compactly supported, nonnegative C^∞ function such that $\psi_t(x) = 1$ for all $x \in \{ty - tz \mid y, z \in A\}$. Lemma 4.2 and the proof of Theorem 4.3 imply that

$$|tA| \le \sup_{\omega\in\mathbb{R}^n} \frac{\widehat{\psi_t}(\omega)}{\widehat{F^n_{p,\alpha}}(\omega)} = \sup_{\omega\in\mathbb{R}^n} \frac{t^n\widehat{\psi}(t\omega)}{\widehat{F^n_{p,\alpha}}(\omega)} = t^n \sup_{\omega\in\mathbb{R}^n} \frac{\widehat{\psi}(\omega)}{\widehat{F^n_{p,\alpha}}(\omega/t)} \le t^n \sup_{\omega\in\mathbb{R}^n} \frac{\widehat{\psi}(\omega)}{\widehat{F^n_{p,\alpha}}(\omega)}.$$ □

The last theorem of this section is a generalization of another result of Leinster [24, Theorem 3.5.6] which complements Theorem 4.4 for subsets with positive volume. Leinster states the result for finite dimensional positive definite normed spaces, but the proof, which will not be repeated here, generalizes immediately from norms to translation invariant, homogeneous metrics on $\mathbb{R}^n$. (*Homogeneous* is used here in the sense that for some $\beta > 0$, $d(tx, 0) = t^\beta d(x, 0)$ for every $x \in \mathbb{R}^n$ and $t > 0$, and not in the sense of possessing a transitive isometry group which follows in any case from translation invariance.) This covers in particular $\left(\ell_p^n\right)^\alpha$ in the entire range $0 < p \le 2$ and $0 < \alpha \le 1$, for which the metric is homogeneous of degree $\alpha \min\{1, p\}$.

Theorem 4.5. *Let d be a positive definite, translation invariant metric on $\mathbb{R}^n$ which is homogeneous of degree $\beta \in (0,1]$, and let $B = \{x \in \mathbb{R}^n \mid d(x,0) \le 1\}$. If $A \subseteq (\mathbb{R}^n, d)$ is compact, then*

$$|A| \ge \frac{\operatorname{vol}(A)}{\Gamma\left(\frac{n}{\beta}+1\right)\operatorname{vol}(B)}.$$

In particular,

$$|tA| \ge \frac{\operatorname{vol}(A)}{\Gamma\left(\frac{n}{\beta}+1\right)\operatorname{vol}(B)} t^n$$

for every $t > 0$.

In the language of [24, Definition 3.4.5], Theorem 4.4 implies that any compact subset $A \subseteq \left(\ell_p^n\right)^\alpha$ has magnitude dimension at most n, and Theorem 4.5 implies that if A also has positive volume, then the magnitude dimension of A is precisely n.

References

[1] Y. Benyamini and J. Lindenstrauss. *Geometric Nonlinear Functional Analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000.

[2] J. Bretagnolle, D. Dacunha-Castelle, and J. Krivine. Lois stables et espaces L^p. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 2:231–259, 1965/1966.

[3] M. M. Deza and M. Laurent. *Geometry of Cuts and Metrics*, volume 15 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1997.

[4] L. E. Dor. Potentials and isometric embeddings in L_1. *Israel J. Math.*, 24(3-4):260–268, 1976.

[5] I. Doust and A. Weston. Enhanced negative type for finite metric trees. *J. Funct. Anal.*, 254(9):2336–2364, 2008.

[6] P. Enflo. On a problem of Smirnov. *Ark. Mat.*, 8:107–109, 1969.

[7] J. Faraut and K. Harzallah. Distances hilbertiennes invariantes sur un espace homogène. *Ann. Inst. Fourier (Grenoble)*, 24(3):xiv, 171–217, 1974.

[8] W. Feller. *An Introduction to Probability Theory and its Applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

[9] M. Gromov. *Metric Structures for Riemannian and Non-Riemannian Spaces*. Modern Birkhäuser Classics. Birkhäuser Boston Inc., Boston, MA, 2007.

[10] P. Hjorth, S. Kokkendorff, and S. Markvorsen. Hyperbolic spaces are of strictly negative type. *Proc. Amer. Math. Soc.*, 130(1):175–181 (electronic), 2002.

[11] P. Hjorth, P. Lisoněk, S. Markvorsen, and C. Thomassen. Finite metric spaces of strictly negative type. *Linear Algebra Appl.*, 270:255–273, 1998.

[12] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[13] I. A. Ibragimov and Yu. V. Linnik. *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff Publishing, Groningen, 1971.

[14] Y. Katznelson. *An Introduction to Harmonic Analysis*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2004.

[15] J. B. Kelly. Metric inequalities and symmetric differences. In *Inequalities, II (Proc. Second Sympos., U.S. Air Force Acad., Colo., 1967)*, pages 193–212. Academic Press, New York, 1970.

[16] J. B. Kelly. Hypermetric spaces and metric transforms. In *Inequalities, III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969)*, pages 149–158. Academic Press, New York, 1972.

[17] S. L. Kokkendorff. Does negative type characterize the round sphere? *Proc. Amer. Math. Soc.*, 135(11):3695–3702 (electronic), 2007.

[18] A. Koldobsky. A Banach subspace of $L_{1/2}$ which does not embed in L_1 (isometric version). *Proc. Amer. Math. Soc.*, 124(1):155160, 1996.

[19] A. Koldobsky. *Fourier Analysis in Convex Geometry*, volume 116 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2005.

[20] S. Kwapień and C. Schütt. Some combinatorial and probabilistic inequalities and their application to Banach space theory. II. *Studia Math.*, 95(2):141–154, 1989.
[21] F. W. Lawvere. Metric spaces, generalized logic, and closed categories. *Rend. Sem. Mat. Fis. Milano*, 43:135–166 (1974), 1973.
[22] T. Leinster. Metric spaces. Post at The n-Category Café, `http://golem.ph.utexas.edu/category/2008/02/metric_spaces.html`, 2008.
[23] T. Leinster. A maximum entropy theorem with applications to the measurement of biodiversity. Preprint, available at `http://arxiv.org/abs/0910.0906`, 2009.
[24] T. Leinster. The magnitude of metric spaces. Preprint, available at `http://arxiv.org/abs/1012.5857`, 2010.
[25] T. Leinster and S. Willerton. On the asymptotic magnitude of subsets of Euclidean space. Preprint, available at `http://arxiv.org/abs/0908.1582`, 2009.
[26] A. J. Lemin. Isometric embedding of ultrametric (non-Archimedean) spaces in Hilbert space and Lebesgue space. In *p-Adic Functional Analysis (Ioannina, 2000)*, volume 222 of *Lecture Notes in Pure and Appl. Math.*, pages 203–218. Dekker, New York, 2001.
[27] C. J. Lennard, A. M. Tonge, and A. Weston. Generalized roundness and negative type. *Michigan Math. J.*, 44(1):37–45, 1997.
[28] J. Lindenstrauss and L. Tzafriri. *Classical Banach Spaces I: Sequence Spaces*, volume 92 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, Berlin, 1977.
[29] K. Menger. Die Metrik des Hilbert-Raumes. *Akad. Wiss. Wien Abh. Math.-Natur. K1*, 65:159–160, 1928.
[30] S. Pavoine, S. Ollier, and D. Pontier. Measuring diversity from dissimilarities with Rao's quadratic entropy: Are any dissimilarities suitable? *Theoretical Population Biology*, 67:231–239, 2005.
[31] I. J. Schoenberg. Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espace distanciés vectoriellement applicable sur l'espace de Hilbert". *Ann. of Math. (2)*, 36(3):724–732, 1935.
[32] I. J. Schoenberg. Metric spaces and completely monotone functions. *Ann. of Math. (2)*, 39(4):811–841, 1938.
[33] I. J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44(3):522–536, 1938.
[34] A. R. Solow and S. Polasky. Measuring biological diversity. *Environmental and Ecological Statistics*, 1:95–107, 1994.
[35] E. M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, Princeton, N.J., 1971. Princeton Mathematical Series, No. 32.
[36] R. S. Varga and R. Nabben. On symmetric ultrametric matrices. In *Numerical Linear Algebra (Kent, OH, 1992)*, pages 193–199. de Gruyter, Berlin, 1993.
[37] C. Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009.
[38] T. Watanabe. Asymptotic estimates of multi-dimensional stable densities and their applications. *Trans. Amer. Math. Soc.*, 359(6):2851–2879 (electronic), 2007.
[39] J. H. Wells and L. R. Williams. *Embeddings and Extensions in Analysis*. Springer-Verlag, New York, 1975. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 84.
[40] H. Wendland. *Scattered Data Approximation*, volume 17 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2005.
[41] S. Willerton. Heuristic and computer calculations for the magnitude of metric spaces. Preprint, available at `http://arxiv.org/abs/0910.5500`, 2009.
[42] S. Willerton. On the magnitude of spheres, surfaces and other homogeneous spaces. Preprint, available at `http://arxiv.org/abs/1005.4041`, 2010.
[43] D. Wolfe. Imbedding a finite metric set in an N-dimensional Minkowski space. *Nederl. Akad. Wetensch. Proc. Ser. A 70=Indag. Math.*, 29:136–140, 1967.

Case Western Reserve University, Cleveland, OH 44106, U.S.A.
E-mail address: `mark.meckes@case.edu`